DOCUMENT RESUME

ED 114 407 TM 004 884

AUTHOR Lewiski, Raymond L.
TITLE Public Policy Concerning the Use of Tests.
NOTE 26p.

EDRS PRICE MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS Civil Rights; Court Cases; *Educational History; Equal Opportunities (Jobs); Ethics; Federal Legislation; Guidelines; Personnel Selection; *Public Policy; Racial Discrimination; Sex Discrimination; Standards; *Testing; *Testing Problems; Test Interpretation; Test Validity

ABSTRACT

The article identifies major events that have impacted on public policy toward testing during the last 20 years. Events were classified, and an interpretation of the interrelationship of events was made. Events were classified as stemming from writings of individual authors, publications of professional organizations, and actions taken by the Legislative, Executive and Judicial Branches of government. The publications of the professional organizations and actions taken by the three branches of government were described in detail. The writings of individual writers were acknowledged but not described. The major concerns of the initiators of events were identified as follows: (a) problems generated by the use of tests for making employment decisions, (b) problems of interpretation generated by semantic differences between the psychological and legal disciplines, (c) problems associated with and resulting from the impact of accountability on the test developer and user. A noticable trend was the shifting of the foci of prime concerns from an emphasis on professional competency to an increasing emphasis on professional integrity. This, in turn, resulted in an increasing resortment to a legalism that holds the test developer and user accountable for their respective product and use. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

PUBLIC POLICY CONCERNING THE USE OF TESTS

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

by
Raymond L. Lewiski
Education Specialist

Extension Course Institute
Air University

ED114407

TM004 884

ABSTRACT

The article identifies major events that have impacted on public policy toward testing during the last twenty years. Events were classified into three catagories and described. Concerns that prompted each event were identified and an interpretation of the interrelationship of events was made. Events were classified as stemming from writings of individual authors, publications of professional organizations, and actions taken by the Legislative, Executive and Judicial branches of government.

The publications of the professional organizations and actions taken by the three branches of government were described in detail. The writings of individual writers were acknowledged but not described.

The major concerns of the initiators of events were identified as follows:

a. Problems generated by the use of tests for making employment decisions.

b. Problems of interpretation generated by semantic differences between the psychological and legal disciplines.

c. Problems associated with and resulting from the impact of accountability on the test developer and user.

A noticable trend was the shifting of the foci of prime concerns from an emphasis on professional competency to an increasing emphasis on professional integrity. This in turn resulted in an increasing resortment to a legalism that holds the test developer and user accountable for their respective product and use.

INTRODUCTION

Public policy concerning the use of tests has been shaped by a multitude of events. Notable events occurring within the past twenty years were reviewed and classified. Events were classified on the basis of who or what institution was instrumental in effecting their occurence. It was found that the most important policy shaping events stemmed from three sources: (1) individual writers, (2) professional organizations, and (3) branches of government.

Among the noteworthy writers in the first classification were Whyte 1956, Baritz 1960, Black 1962, Gross 1962, Hoffman 1962, Fincher 1964 and 1973, Enneis 1964-1969, Ash 1966, Krug 1966, and Lopez 1966. Three noteworthy professional organizations were the American Psychological Association (APA), the American Educational Research Association AERA), and the National Council on Measurement in Education (NCME). Noteworthy events from the third classification consisted of executive orders, legislative acts, and court rulings. Tables 1, 2, and 3 chronologically list noteworthy events classified as originating from individual writers, professional organizations, and branches of government.

Events such as noteworthy publications by individual writers were acknowledged in Table I but not described in detail. Noteworthy events that stemmed from professional organizations and branches of government were given more consideration. Each event classified in these two groupings was described, and an interpretation of its relationship to other events was given. Also an attempt was made to identify major trends, issues, and concerns for each event. The first of these events was the publication of "Ethical Standards of Psychologists" by the American Psychological Association.

Ethical Standards of Psychologists

In 1953 the American Psychological Association (APA) published its "Ethical Standards of Psychologists." This publication was one of the first attempts by an association to initiate a formal set of standards governing the use of tests. During the next two decades, a number of other noteworthy publications that pertained to tests were published. Among these were "Technical Recommendations for Psychological Tests and Diagnostic Techniques" in 1954, "Technical Recommendations for Achievement Tests" in 1955, "Standards for Educational and Psychological Tests and Manuals" in 1966, and "Standards for Educational and Psychological Tests" in 1974. In each of these publications, the dominant theme was interpreted as representing the issues and concerns of the time.

One of the major concerns of the association during the early fifties was the prevention test abuses. It was thought that prevention of test abuses could be accomplished by insuring that test constructors and users were properly trained. The prevailing attitude was that professionally developed tests administered by competent persons would keep test abuses at a minimum. Little emphasis was placed on issues concerning the professional integrity of the test constructors or test users: The focus of attention was on their professional competencies. Thus, in keeping with this belief, a major topic in the 1953 publication pertained to the background knowledge required of persons using various types of test.

Tests were classified into three levels: A, B, and C. Examples of tests classified at each level were given and competencies expected of persons who used tests classified at the various levels were described.

Level A tests consisted of those tests that could be administered, scored and interpreted with the aid of a manual by a person who had only a general orientation to the kind of institution or organization for which the test was being administered. Achievement or proficiency tests were examples of tests at this level.

Level B tests required that the tester have some technical knowledge of test construction and use. The tester was required to have competency in supporting psychological and educational fields such as statistics, individual differences, psychology of adjustment, personnel psychology and guidance. Examples of level B tests would include aptitude tests and adjustment inventories that apply to normal population.

Level C tests consisted of tests and aids, and required that the tester have a substantial understanding of testing and supporting psychological fields, together with supervised experience in the use of these devices. Projective tests and individual mental tests would be included at this level.

The major objectives of the 1953 publication were the classification of tests into levels and the establishment of minimum training standards. These objectives were in keeping with the belief that professionally developed tests administered by competent persons would keep tests abuses at a minimum. The possibility of deliberate test abuses and the legal implications that would result from such abuses was not a major concern at this time. The prevailing belief was that test abuses resulted more from ignorance than intent.

Perhaps the greates influence on public policy toward testing resulting from the "Ethical Standards of Psychologists"

was its role as a ferment of discussion. The need for a set of guidelines governing the use of tests was recognized by a number of professional associations. Among these were the American Education Research Association and The National Council on Measurements used in Education. The APA publication fermented discussion within these associations and was the beginning of a joint effort between associations to establish guidelines governing testing.

Standard for Educational and Psychological Tests and Manuals

The "Standards for Educational and Psychological Tests and Manuals" was a publication which resulted from the efforts of three associations. The sequence of events leading to this publication was as follows: In March 1954, the American Psychological Association (APA) issued its "Technical Recommendations for Psychological Tests and Diagnostic Techniques." In January 1955, the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME) jointly issued their "Technical Recommendations for Achievement Tests." In the latter part of 1963 a joint APA-AERA-NCME committee of eight members was constituted for the purpose of formulating a set of technical standards accepted by each association. Three years later in 1966 the associations agreed on and jointly published their "Standards for Educational and Psychological Tests and Manuals." This publication was a major contribution toward the formation of a set of standards and recommendations for persons who use educational and psychological measurements.

The members of the three associations who were the architects of the standards did not restrict their efforts solely to the issues of testing technology; they were also concerned about the human aspects of testing. The authors of the 1966 standards described the impact of tests on the individuals tested as follows:

> Psychological and Educational tests are used in arriving at decisions which may have great influence on the ultimate welfare of the persons tested, on educational points of view and practices, and on development and utilization of human resources. Test users, therefore, need to apply high standards of professional judgement in selecting and interpreting tests, and test producers are under obligation to produce tests which can be of the greatest service. The test producer, in particular, has the task of providing sufficient information about each test so that the users will know what reliance can safely be placed on it.[1]

This quotation depicts the attitude that must have prevailed among the developers of the 1966 Standards. The concern about what reliance can safely be placed on a test was of particular note because this concern was to increase in intensity and eventually generate a host of moral and legal issues about test validity. Basically the foci of concerns had shifted from an emphasis on the test constructor's and user's competency to an emphasis on what reliance can safely be placed on a test.

Six topics pertaining to tests were covered in the 1966 Standards: (1) Dissemination of information, (2) Interpreting, (3) Validity, (4) Reliability, (5) Administration and scoring, and (6) Scales and norms. Of these topics, the one that generated the most controversy was the one on validity. The coverage given to validity was criticized as being both imprecise and insufficient. The decriptions of the various types of validity given in the 1966 Standards may have been precise and sufficient for a person schooled in psychometrics but for other readers they were vague and incomplete. Three types of validity were described; namely: Content Validity, criterion-related validity, and construct validity. Future guidelines from the Equal Opportunity Commission were to recommend each of th se types of validity for meeting minimum standards for ter. validation.

The Motorola Case

The case of Motorola, Inc. Vs. Illinois Fair Employment Practices Commission and Leon Myart was the first major court case pertaining to the use of tests for employment purposes. The main events of this case beginning 15 July 1963 and ending 24 March 1966 were outlined by Wallace, Kissinger and Reynolds in Appendix A of the article "Testing of Minority Group Applicants for Employment." The following was based on their chronology of the Motorola Case.

On 15 July 1963 Leon Myart, a black, applied for a job as a television phaser and analyzer at the Franklin Plant of Motorola, Inc., in Franklin, Illinois. He was required to take a five-minute intellignece test administered by the company and was interviewed. He was not offered a job, nor was he informed whether he qualified for the job.

On 29 July 1963 Myart filed a complaint with the Illinois Fair Employment Practices Commission and the President's Committee on Equal Employment Opportunity alleging that his not being hired was due to racial discrimination. The case was examined by Robert Bryant of the Illinois Fair Employment Practices Commission on 27 and 28 January 1964, and the

question whether a general ability test could be used as a basis for selection for employment was raised. Upon examining the case, Bryant ruled that the intellignece test, General Ability Test no. 10, could not be used as a basis for disqualifying Myart for employment because the test had been normed on "advantage groups" and did not "lend itself to equal employment opportunity to qualify for the hitherto culturally deprived and disadvantaged groups." He therefore directed Motorola to offer Myart a job and to discontinue using test no. 10.

The decision was appealed and the case was revieded before the full commission four times during 1964. Reviews of the case were held on April 18, May 25, July 14-15, and November 18.

On November 18 a unanimous decision was issued by the commission. The members of the commission agreed that Myart had been denied employment because of his race, but they did not direct that he be hired by the company. They did direct however, that he be compensated one-thousand dollars by Motorola, Inc.

The case was appealed to the Illinois Circuit Court. The circuit court reversed the decision that Myart be paid one-thousand dollars compensation; however, the findings on discrimination were upheld.

The case was then appealed to the Illinois Supreme Court. On March 24, 1966, the Illinois Supreme Court reversed the judgement of the circuit court on the grounds that the alleged unfair employment practice was not established by a preponderance of evidence.

This ended a case that had been under review for over two and one-half years, a case that had generated controversy and undergone a number of reverse decisions, a case that had received national publicity and, finally, a case that was a landmark in judicial involvement in public policy toward testing. The importance of the case was not attributed to the final ruling of the Illinois Supreme Court. The importance of the case was due to the publicity that was given to the directives issued by the Illinois Fair Employment Commission. The directives issued by the commission on January 1964 gained national attention. The impact of the publicity was in turn felt in congress where civil rights legislation was being debated. This in turn prompted Senator Tower, a Republican from Texas, to attach an amendment to the Civil Rights Act of 1964.

The Tower Amendment to The Civil Rights Act of 1964

During 1964, civil rights legislation was being debated in Congress. The topics of debate covered a wide spectrum, including public education, public facilities, housing, employment practices, etc. The question of unlawful employment practices was raised and debated. During debate of this issue, an amendment was introduced by Senator Tower of Texas. This amendment later became the Tower Amendment to Title VII of the Civil Rights Act of 1964.

The Tower Amendment was prompted by the Illinois Fair Employment Practice Commission's ruling on the Motorola Case. The commission had disbarred the use of general ability tests by employers, if said tests were to be used as a basis for selecting persons to be hired. The Tower Amendment had the effect of establishing, in legal terms, the right of an employer to use "professionally developed ability tests" for the purpose of selecting persons for employment provided that the tests were not intentionally discriminatory. Thereby, the amendment represented a major attempt by the legislative branch of government to affect public policy governing the use of tests.

The Tower Amendment was embodied in Section 703 (h) of Title VII of the Civil Rights Acts of 1964. It reads as follows:

> Nor shall it be an unlawful employment practice for an employer to give and act upon the results of any professionally developed ability test provided that such tests, its administration or action upon the results is not designed, intended or used to discriminate because of race, color, religion, sex, or national orgin.[2]

The wording of the amendment led to a controversy over the use of and meaning of the term "ability test". The question was raised whether a legal interpretation of the term "ability test" was to include broad general abilities that tend to measure the person in the abstract or be restricted to specific job related abilities that could be empirically verified. In other words, were mental ability tests or intellignece tests also to be classed as "ability tests?" Members from both the psychological and legal professions were in disagreement over how the term "ability test" was to be interpreted. An answer to this question was not easily forthcoming.

The Emergence of Executive Influence on Public Policy Toward Testing Executive Order No. 11246

On 24 September 1965 President Johnson issued Executive Order 11246 entitled "Equal Employment Opportunity". This order superseded Executive Order No. 10590 (19 June 1955), 10722 (5 August 1957), 10925 (6 March 1961), 11114 (22 June 1963) and 11162 (28 June 1964).

The order consisted of four parts, three of which indirectly pertained to public policy toward the use of tests. These three were (1) nondiscrimination in government employment, (2) nondiscrimination in employment by government contractors and subcontractors, and (3) nondiscrimination provisions in federally assisted construction contracts.

Part I consisted of policy statements by the executive branch of government on discrimination in federal employment. In general, a policy of equal opportunity was to apply to every aspect of federal employment. Discrimination on the basis of race, creed, color, or national orgin was specifically prohibited. Each executive department and agency was required to use an existing staff or establish an office for the purpose of carrying out these policies. For example, the task of insuring equal employment opportunity programs for civilian employees was assigned to the existing staff of the Civil Service Commission. Other governmental departments, such as the Department of Labor, had to establish an office to carry out such parallel functions.

Parts II and III dealt with discrimination as it related to employment practices by government contractors, subcontractors and federally assigned construction contractors. The topics covered in Part II were (A) duties of the Secretary of Labor, (B) contractors' agreements, (C) powers and duties of the Secretary of Labor and the contracting agencies, (D) sanctions and penalties, and (E) certificates of merit; whereas, Part III dealt with nondiscrimination provisions in federally assisted construction contracts.

The issuance of Executive Order 11246 set the stage for an eventual impact on public policy toward testing. One feature of the order was a requirement that nondiscrimination clauses be included in all government contracts, sub-contracts, and federally assisted construction contracts. The Secretary of Labor was charged with the responsibility for insuring that contractors comply with such clauses. To carry out this obligation, the Secretary of Labor established The Office of Federal Contract Compliance, Equal Employment Opportunity, Department of Labor. In turn, authority was delegated to

this office for carrying out its responsibilities. One responsibility of this office was to develop and issue guidelines governing employment practices. Thus, authority had been delegated and the stage set for an eventual impact by an agency of the executive branch of government on public policy toward the use of tests.

The Executive Branch of Government Acknowledges Sex as a Possible Basis for Discrimination (Executive Order 11375)

A noticeable phenomenon about the text of Executive Order 11246 was the exclusion of sex as a possible basis for discrimination. The exclusion of sex was particularly noticeable in view of the fact that Title VII of the Civil Rights Act of 1964 dealt with the problem of Equal Employment Opportunity, and Section 703 of Title VII was entitled "Discrimination because of Race, Color, Religion, Sex or National Orgin;" whereas, Executive Order 11246 entitled "Equal Employment Opportunity" issued almost a year after the passage of the 1964 Civil Rights Act excluded sex as a possible basis for discrimination. Order 11246 acknowledged race, creed, color and national orgin as possible basis for discrimination but not sex. This remained the official position of the executive branch of government until 1967 when Executive Order 11375 was issued. In effect, Order 11375 amended 11246 to include sex as a possible basis for discrimination. In each case "Discrimination because of race, creed, or national orgin" was amended to read "Discrimination because of race, color, religion, sex or national orgin."

The Equal Employment Opportunity Commission's "Guidelines on Equal Employment Testing Procedures"

On 2 July 1964 civil rights legislation was passed by Congress that would eventually have a considerable effect on public policy toward the use of tests. A year later on 2 July 1965 the act became effective. One intent of the act was to insure that all Americans would be considered for hiring and promotion on the basis of their ability and qualifications, without regard to race, color, religion, sex or national orgin. To insure that this intention would be carried out, the act provided for the establishment of the Equal Employment Opportunity Commission (EEOC).

A prime responsibility of the EEOC was to conduct technical studies in support of the purposes of Title VII. It soon became apparent that the purposes of Title VII could be interpreted in more than one way; of particular note were the different interpretations given for Section 703 (h), the

Tower Amendment. A major dispute developed over how this section was to be interpreted. To resolve this dispute, the EEOC formulated a set of guidelines which were published in the Federal Register on 24 August 1966. The publication was entitled "Guidelines on Equal Employment Testing Procedures." This publication was the first set of guidelines to be issued by the EEOC, and they were to be met with both success and failure.

From its very beginning, the EEOC was embroiled in controversies--some semantical, others attitudinal. From the semantical side, the wording of the Tower Amendment had led to controversy over what constituted an ability test, while on the attitudinal side there was a noticeable shift of the foci of prime concerns from an emphasis on professional competency to an emphasis on professional integrity.

The semantical problem has been viewed as resulting from a communications gap between two distinct disciplines, namely those of psychometrics and law. In 1969 Enneis perceived the problem as follows:

> During the four years that the Equal Employment Opportunity Commission has been in operation, the issues of employment testing have been among the more persistent and difficult ones. As far as the EEOC is concerned, it is probably correct to say that psychological testing involves the most direct confrontation of a scientific discipline with legal definitions of employment discrimination.[3]

The resolution of semantic differences between the psychological and legal disciplines was to be the major concern of forthcoming guidelines by the EEOC.

From the standpoint of attitude, the shift in concerns meant that issues would no longer focus merely on the competency or incompetency of the test constructor and user; it would also focus on their intentions and professional integrity. The following quote from the 1966 EEOC guidelines exemplifies both the emerging semantical issues and shifting of concerns.

> The commission accordingly interprets professionally developed ability tests to mean a test which fairly measures the knowledge or skill required by the particular job or class of jobs which the applicant seeks, or which fairly offers the employer a chance to measure the applicant's ability to perform a particular job or class of jobs. The fact that a test was prepared by an individual or organization claiming expertise in test

preparation does not, without more, justify its use within the meaning of Title VII.[4]

In brief, the 1966 guidelines were vaguely written and tended to raise issues, not resolve them. The replacement of "professionally developed ability tests" by "a test which fairly measures knowledge or skills and fairly offers the employer a chance to measure the applicant's ability" did little toward resolving the controversy. The guidelines did not indicate what a fair measure would be or how fairness would be judged. Also, what constituted "more" in "without more" in the above quotation was not clearly defined; at best, it alluded to an emerging attitude that test developers and users ought to be accountable for their respective products and uses.

Despite the acknowledged shortcomings of the 1966 EEOC Guidelines, their publication was a noteworthy event. They were the first set of guidelines to be issued in compliance with the 1964 Civil Rights Act and represented a major influence, stemming from the legislative branch of government, on public policy toward testing. They were to remain in effect until 1 August 1970, at which time they would be superseded by a new set of guidelines entitled "Guidelines on Employee Selection Procedures."

The Equal Employment Opportunity Commissions "Guidelines on Employee Selection Procedures"

On the first of August 1970 a new set of guidelines entitled "Guidelines on Employee Selection Procedures" was issued by the EEOC. The new guidelines superseded and enlarged upon the 1966 guidelines. A noticeable trend in the new guideline was the tendency to describe in more detail the topics discussed and terminology used. Two terms of particular importance were "test" and "discrimination."

The term "test" was defined as any paper and pencil or performance measure used as a basis for any employment decision. Paper and pencil or performance measures included measures of general intelligence; mental ability and learning ability; specific intellectual abilities; mechanical, clerical and other aptitudes; dexterity and coordination; knowledge and proficiency; occupational and other interests; and attitudes, personality or temperament. Also included were all formal, scored quantified or standardized techniques of assessing job suitability such as specific qualifying or disqualifying personal history or background requirements, specific educational or work history requirements, scored interviews, biographical information blanks, interviews rating scales, scored application forms, etc. Any of the above

measures, when used as a basis for making employment decisions, constituted a "test", whereas employment decisions included decisions affecting one's eligibility for hire, transfer, promotion, membership, training, referral or retention.

Discrimination was defined as the use of any test which adversely affects hiring, promotion, transfer or any other employment or membership opportunity of classes protected by Title VII of the Civil Rights Act of 1964 unless: (a) the test has been validated and evidences a high degree of utility and (b) the person giving or acting upon the results of the particular test can demonstrate that alternative suitable hiring, transfer or promotion procedures are unavailable for his use.

Topics on validity were discussed in detail. Among these were topics on Evidence of Validity, Minimum Standards for Validation, Presentation of Validity Evidence, Use of Other Validity Studies and Assumption of Validity.

The guidelines disallowed any assumption of validity and placed the burden of proof of validity on the tester. Under no circumstances could the general reputation of a test, its author or its publisher, or casual reports of a test's utility be accepted in lieu of evidence of validity. In every case in which a test was used for making an employment decision, evidence of the test's validity was to be supplied by the tester.

Validation requirements varied with each testing situation. Under certain conditions, content or construct validity were acceptable for meeting validity requirements. For example, evidence of content validity alone would be acceptable provided that the test were well developed and consisted of suitable samples of the essential knowledge, skills or behaviors composing the job in question. However, when technically feasible, a criterion-related validity was required and could not be substituted for by content or construct validities. The state of being "technically feasible" was described in the guidelines as follows:

> The term "technically feasible" as used in these guidelines means having or obtaining a sufficient number of minority individuals to achieve findings of statistical and practical significance, the opportunity to obtain unbiased job performance criteria, etc.[5]

Persons claiming the absence of technical feasibility were required to demonstrate positive evidence of this absence. In other words, evidence of content (in the case of job knowledge or proficiency tests) or construct (in the

case of trait measures) validity, would be acceptable only if it could be positively demonstrated that a criterion-related validity was technically unfeasible.

The "Standards for Educational and Psychological Tests and Manuals" published by the American Psychological Association (APA) were specifically referred to in the topic on "Minimum Standards for Validation." Evidence of content, construct or criterion-related validity called for in the guidelines was to be subjected to the same standards for those types of validity as described in the APA publication. For example, if content validity were allowed as evidence of test validity, then the standards governing content validity in the APA publication would be applicable to the tester. This included such requirements as obtaining an adequate sample, describing the credentials of the experts who selected the items, and reporting the extent of agreement between independent judgements about each item.

Another type of validity referred to in the guidelines but not mentioned in the APA publications was differential validity. Differential validity was included as a validity requirement in the 1970 guidelines. A precise definition of differential validity was not offered; however, various aspects that pertained to the notion of differential validity were described. In brief, to meet the differential validity requirements, data must be generated and results separately reported for minority and non minority groups whenever technically feasible. Here again the question of technical feasibility was raised and the burden of demonstrating technical unfeasibility fell on the tester.

The 1970 guidelines, like the 1966 guidelines, clarified some questions and generated others. Detailed descriptions of terms such as "test" and "discrimination" tended toward clarification; however, vaguely defined phrases such as "differential validity" and "technically feasible" insured that semantic differences between the legal and psychological disciplines would remain.

Despite these shortcomings in the intended goal of clarifying meaning, the 1970 guidelines met with some success and marked a high point in legislative influence on public policy toward testing. The use of tests for employment decisions had increased considerably since the passage of the 1964 Civil Rights Act, and the number of possible misuses of tests had likewise increased. Thus, the 1970 guidelines provided a much needed interpretation of the intent of Title VII of the 1964 act and, thereby, contributed to the implementation of nondiscriminatory personnel policies. Then too, in all likelihood these guidelines abetted the ferment for a

court test of Title VII. Such a court test was soon to be forthcoming in Griggs Vs. Duke Power Company.

Griggs Vs. Duke Power Company

The case of Willie S. Griggs et al., Petitioners., Vs. Duke Power Company was the second major court case pertaining to public policy toward the use of tests. The first case, Myart Vs. Motorola, was decided in the Illinois Supreme Court on March 1966. The Griggs Vs. Duke Power Company case was decided in the U.S. Supreme Court on the 8th of March 1971. The former case differed from the latter case in both the level of the deciding court and the direction of the decision rendered. Unlike the former case, the Griggs Vs. Duke Power Company case was a U.S. Supreme Court decision and in favor of the original petitioners. Consequently it has had a greater effect on public policy toward the use of tests than the Motorola case.

The Griggs Vs. Duke Power Company case began as a class action suit by black employees at Duke Power Company's Dan River steam station located at Draper, North Carolina. The station had 95 employees, 14 of whom were black. Thirteen of the 14 black employees were petitioners in the suit. The plantiffs claimed that the company's employment practices violated their civil rights. Specifically challedged was the company's employment requirement that any employee or potential employee have a high school education or pass a standardized general intelligence test as a condition for transfer to a new job (in the case of an old employee) or initial employment (in the case of a new employee). The petitioners claimed that these requirements for hire and transfer operated to render ineligible a disproportionate number of blacks.

The case was brought before the United States District Court for the middle district of North Carolina at Greensboro, North Carolina. The District Court dismissed the complaint. The dismissal was appealed and the appeals court remanded that in absence of a discriminatory purpose, requirement of a high school education or the passing of a standardized general intelligence test as a condition of employment in or transfer to jobs was permitted by the Civil Rights Act. The claim was rejected that such requirements operated to render ineligible a markedly disproportionate number of Negroes and were therefore unlawful under the act unless shown to be job-related. Certiorari was granted and the case was sent to the Supreme Court for review. The opinion of the court was delivered by Chief Justice Burger. The Supreme Court reversed the appeals court's remand, holding that the employer was prohibited by

provisions of the civil rights act from requiring a high school education or passing of standardized general intelligence where (a) neither standard was shown to be significantly related to successful job performance, (b) both requirements operated to disqualify Negroes at a substantially higher rate than white applicants, and (c) jobs in question formerly had been filled only by white employees as part of a longstanding practice of preference to whites. This decision was to have a profound impact on public policy toward testing. Prior to the Supreme Court's ruling, the EEOC's "Guidelines on Employee Selection Procedure" offered an interpretation of the intent of the Civil Rights Act of 1964 and recommended necessary measures an employer should take to be in compliance with the act's intent. By contrast, the Supreme Court's decision was a definitive statement of the act's intent and, henceforth, tests used for the purpose of making employment decisions were required by law to be job-related, and the burden of proving job-relatedness was placed on the employer.

Justice Burger interpreted Congressional intent as follows:

> Nothing in the act precludes the use of testing or measuring procedures; obviously they are useful. What Congress has forbidden is giving these devices and mechanisms controlling force unless they are demonstrably a reasonable measure of job performance. Congress has not recommended that the less qualified be preferred over better qualified simply because of minority orgins. Far from disparaging job qualifications as such, Congress has made such qualifications the controlling factor, so that race, religion, nationality, and sex become irrelevant. What Congress has commanded is that any tests used must measure the person for the job and not the person in the abstract.[6]

Basically the foci of concerns had shifted from emphasis on the test constructor's and user's competency to an emphasis on what reliance could safely be placed on a test.

The Supreme Court's decision rendered a definitive statement governing some aspects of testing and employment decision making; however, other aspects were not considered in this case. For example, no ruling was rendered about the legal status of testing requirements that take into account capabilities for the next succeeding position. For instance, in a context in which a potential for future promotion could be shown to be a genuine business need, a test that was shown to be job-related to succeeding positions might be considered as a bonified test. At this time the status of such a test has not been determined.

Part 60-3 -- Employee Testing and Other Selection Procedures

Less than two months after the Griggs Vs. Duke Power Company case was decided, a notice of proposed rule making for amending Chapter 60 of Title 41 was published in the Federal Register. The usual 30 days were given to interested persons to submit written comments, suggestions, or objections regarding the proposed amendments and finally, on 2nd October 1971, an amendment to chapter 60, "Part 60-3 Employee Testing and Other Selection Procedures" was published in the Federal Register and became effective on that day.

The guidelines given in "Part 60-3 Employee Testing and Other Selection Procedures" were basically the same as the "Guidelines on Employee Selection Procedures" issued on the 1st of August 1970 by the EEOC. Both sets of guidelines were intended to impose the same basic requirements on employers, contractors, or other persons covered by them. The guidelines differed only in the language arising from the different legal authority of the two agencies. This was noteworthy in that it was one of the first deliberate attempts to arrive at a concensus between governmental agencies toward the question of what should be contained in their guidelines. The question of what should be included in a uniform set of guidelines applicable to all governmental agencies was to become a major concern for the EEOC.

The Equal Employment Opportunity Act of 1972

On March 24, 1972, The Equal Employment Opportunity Act (Public Law 92-261) was approved. The enactment of this law was a major step forward in assuring an equal employment opportunity to minority groups and women in our society. In turn, it was a major event affecting public policy toward the use of tests. The '72 act amended Title VII of the Civil Rights Act of 1964, expanding its coverage and adding enforcement powers.

The duties and responsibilities of the EEOC were broadened, giving it quasi-legal enforcement powers. Section 706 (a) as amended empowered the commission to prevent any person from engaging in any unlawful employment practices as set forth in Section 703 (Discrimination Because of Race, Color, Religion, Sex, or National Orgin) and Section 704 (Other Unlawful Employment Practices). Cases involving a government, governmental agency, or political subdivision were charged to the Attorney General. Section 707 (c) as amended provided for the commission's assumption of the functions charged to the Attorney General and his office. These functions were to be assumed within two years after the date of enactment of the Equal Employment Opportunity Act of 1972.

Another amendment that will most likely have a consierable effect on public policy toward the use of tests was Sec. 715. This section provided for the establishment of the Equal Employment Opportunity Coordination Council (EEOCC) and reads as follows:

> There shall be established an Equal Employment Opportunity Coordinating Council (hereinafter referred to in this section as the Council) composed of the Secretary of Labor, the Chairman of the Equal Employment Opportunity Commission, the Attorney General, the Chairman of the United States Civil Service Commission, and the Chairman of the United States Civil Rights Commission, or their respective delegates. The Council shall have the responsibility for developing and implementing agreements, policies and practices designed to maximize effort, promote efficiency, and eliminate conflict, competition, duplication and inconsistency among the operations, functions and jurisdictions of the various departments, agencies and branches of the Federal government responsible for the implementation and enforcement of equal employment opportunity legislation, orders, and policies. On or before July 1 of each year, the Council shall transmit to the President and the Congress a report of its activities, together with such recommendations for legislative or administrative changes as it concludes are desirable to further promote the purposes of this section.[7]

The Council has issued a discussion draft of its pending "Uniform Guidelines on Employee Selection Procedures." These should be approved sometime this year (1974).

Standards for Educational and Psychological Tests

The "Standards for Educational and Psychological Tests" was published in 1974 by the American Psychological Association (APA). The publication was the joint work of many psychologists and educators. The principal author was Professor Robert M. Guion, Bowling Green State University, Bowling Green, Ohio. Basically, the publication was a revision of the 1966 "Standards for Educational and Psychological Tests and Manuals." Like the 1966 publication, the 1974 publication was a joint effort by the American Psychological Association, the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME).

The publication of the 1974 Standards was prompted by a number of events. First, the guidelines issued by the Equal Employment Opportunity Commission in August 1970 made reference to the 1966 Standards. This focused attention on the

standards which were lacking in clarity on some topics and did not adequately cover others. Then on 8 March 1971 the historic Supreme Court decision of Griggs Vs. Duke Power Company impacted further on public policy toward the use of tests. This decision greatly increased the number of discrimination suits and in turn increased the need for a new set of standards. Shortly thereafter, on 2 October 1971, the Office of Federal Contract Compliance, Equal Employment Opportunity Department of Labor, issued a set of guidelines by authority of Executive Order 11246. Again specific reference was made to the 1966 APA Standards. By this time it was apparent that the 1966 Standards needed to be revised. Finally, as plans for revision were underway the Equal Employment Opportunity Act of 1972, Public Law (92-261) further influenced public policy toward the use of tests. The 1972 Act broadened the coverage of Title VII of the 1964 Act and included employment decisions by educational instituitions within its jurisdiction. Each of these events had its impact on the 1974 Standards.

Another factor impacting on the 1974 Standards was the prime concerns of its authors. The prime concerns of the authors of the 1974 Standards differed somewhat from the concerns expressed by the authors of the 1966 Standards. The authors of both publications shared similar concerns about the human aspect of testing. However, changing conditions dictated some new concerns. The following quote from the 1971 Standards reflects some of these concerns:

> Part of the stimulus for revision is awakened concern about problems like invasion of privacy of discrination against members of groups such as minorities or women. Serious misuses of tests include, for example, labeling Spanish-speaking children as mentally retarded on the basis of scores on tests standardized on a "representative sample of American children," or using a test with a major loading on verbal comprehension without appropriate validation in an attempt to screen out large numbers of blacks from manipulative jobs requiring minimal verbal communication.[8]

The examples cited above were among the concerns shared by the authors of 1974 Standards.

Another factor impacting on the 1974 Standards was an emerging notion of accountability. Accountability has become a concern for most professional groups; ie, teachers, administrators, politicians, etc. Accountability as it applies to public policy toward testing requires that both the test constructor and test user be held responsible for their respective product or use. The requirement that they

be responsible raises some complex questions, and poses problems that have both ethical and legal ramifications. Questions such as "how would responsibility be enforced" and "what conditions were necessary before responsibility could be actualized" were among those raised. The mere delegation of responsibility does not insure accountability. To actualize accountability, two things must be done. First, some person or agency must scrutinize the test developer and tester and second, the prerequisite conditions to responsibility must be identified and monitored. Two such prerequisites to responsibility have been identified. These were (1) that the test developer and tester be competent and (2) that they be fair. The inclusion of fairness as a prerequisite to responsibility meant that the scrutiny of the individual in question would not be limited solely to his professional competency but would also be directed toward his professional integrity. This represented a major shift from the previously held belief that all that was necessary to prevent test abuses was to insure that persons involved in testing were competent.

The authors of the 1974 Standards were sensitive to the concerns and issues that resulted from using tests for making employment decisions. This was attested to by the numerous examples pertaining to employment practices cited in the publication. They were also sensitive to the semantic differences that existed between the psychological and legal disciplines. This was attested to by the coverage given to the topic on validity.

Validity has been a particularly troublesome concept. The topic on validity in the 1966 Standards was lacking in both coverage and clarity. From a psychometric point of view the 1966 descriptions may have been acceptable, but from a legal point of view they were inadequate. In view of the deficiency, a greater emphasis was placed on the topic of validity in the 1974 Standards and examples cited for content, construct, criterion related, and differential validities were more definitive. It is important to note that the standards were only recommendations and had no legal status per se; however, the issues reviewed pertaining to enforcement, competency and fairness were also vital legal issues. Thus, the 1974 Standards will most likely influence forthcoming legal opinions and may resolve many of the semantic differences that exist between the psychological and legal disciplines.

Some issues have not been resolved. For example, the relationship of validity to the notion of job relatedness has been intrepreted differently by different persons. Guion[9] suggested that an employment test may provide a basis for inferences that have criterion-related validity, or construct validity, or content validity, or all of these, and still not

be job related, Guion was unwilling to equate "validity" and "job relatedness." Also, there is no general agreement on the prior question of what is fairness and how it can be measured. Various math models have been suggested for assessing fairness. Cleary[10] (1968) suggested that a regression analysis be used as a basis for determining and measuring fairness. Thorndike[11] (1971) recommended that a quota system be used and Darlington[12] (1971) approached the problem in terms of a partial r, correlational technique. As of now the issues surrounding the notion of job relatedness and fairness have not been resolved. Possible, these issues will be clarified in the forthcoming "Uniform Guidelines on Employer Selection Procedures."

TABLE I

A CHRONOLOGICAL LIST OF ARTICLES THAT HAVE HAD A NOTEWORTHY EFFECT ON PUBLIC POLICY TOWARD THE USE OF TESTS

Date	Author	Article
1956	Whyte, W. H., Jr.	The Organization Man. Garden City, N.Y., Doubleday , 1956
1960	Baritz, L.,	The Servants of Power, Middletown, Conn,; Wesleyan University Press, 1960.
1962	Gross, M. L.,	The Brain Watchers. New York; Random House, 1962.
1962	Hoffman, B.,	The Tyranny of Testing. New York, Crowell-Collier, 1962.
1963	Black, H.,	They Shall Not Pass. New York,Morrow, 1963.
1964	Fincher, C.,	Testing Critics and Criticism. Atlanta Economic Review, 1964, 3-7.
1966	Ash, Philip,	The Implications of the Civil Rights Act of 1964 for Psychological Assessment in Industry. American Psychologist, 1966, 21, 797-703.
1966	Krug, R. E.,	Some Suggested Approaches for Test Development and Measurement. Personnel Psychology, 19, No. 1, 24-35.
1966	Lopez, Felix, Jr.,	Current Problems in Test Performance of Job Applicants. Personnel Psychology, 19, No. 1, 10-18.
1967	Enneis, William H.,	Statement Before House Post Office and Civil Service Subcommittee.*
1967	" " "	Discrimination: Planned and Accidental.*
1969	" " "	Personnel Testing and Equal Employment Opportunity.*
1969	" " "	Misuses of Tests.*
1969	" " "	Minority Employment Barriers from the EEOC Viewpoint.*
1970	" " "	Statement on Personnel Testing and Selection.*
1970	" " "	Use of Nontest Variables in the Government Employment Setting.*

* Equal Employment Opportunity Commission, Personnel Testing and Equal Employment Opportunity, U.S. Government Printing Office, Washington, D.C. 20402, Dec, 1970.

TABLE II

A CHRONOLOGICAL LIST OF NOTEWORTHY EVENTS INITIATED BY PROFESSIONAL ORGANIZATIONS THAT HAVE AFFECTED PUBLIC POLICY TOWARD TESTING

	EVENT	SOURCE	DATE
1.	Publication of "Ethical Standards of Psychologists."	American Psychological Association (APA)	1953
2.	Publication of "Technical Recommendations for Psychological Tests and Diagnostic Techniques."	American Psychological Association (APA)	March 1954
3.	Publication of "Technical Recommendations for Achievement Tests."	American Educational Research Association (AERA) and The National Council on Measurement in Education. (NCME)	Jan 1955
4.	Publication of "Standards for Educational and Psychological Tests and Manuals."	APA, AERA, and NCME	1966
5.	Publication of "Standards For Educational and Psychological Tests."	APA, AERA, and NCME	1974

TABLE III

A CHRONOLOGICAL LIST OF EXECUTIVE, LEGISLATIVE AND JUDICIAL EVENTS THAT HAVE SHAPED PUBLIC POLICY TOWARD TESTING

EVENT	SOURCE	DATE
1. Myart vs. Motorola Co.	Judicial	15 July 1963 - 24 March 1966
2. The Civil Rights Act of 1964	Legislative P.L. 88-352	2 July 1964
3. Issuance of Executive Order on Equal Employment	Executive Order 11246	28 September 1965
4. Publication of EEOC's "Guidelines on Equal Employment Testing Procedures"	Legislative 88-352	24 August 1966
5. Amendment of Executive Order 11246	Executive Order 11375	17 October 1967
6. Publication of EEOC's "Guidelines on Employee Selection Procedures"	Legislative 88-352	1 August 1970
7. Willie S. Griggs et al, Petitioners, Vs. Duke Power Company	Judicial 401 U.S. 124	8 March 1971
8. Publication of "Part 60-3 Employee Testing and other Procedures" by Office of Federal Contract Compliance, Equal Employment Opportunity, Department of Labor	Executive Order 11246	2 October 1971
9. The Equal Employment Opportunity Act of 1972	Legislative, Public Law (92-261)	March 1972

FOOTNOTES WITH REFERENCES

1. American Psychological Association, American Educational Research Association, and National Council on Measurement in Education (joint committee). Standards for Educational and Psychological Tests and Manuals. Washington, D.C.; American Psychological Association, Inc. 1966. page 1.
2. Public Law 88-352, as approved by the President on July 2, 1964; Civil Rights Act of 1964 with Explanation Commerce clearing House, Inc., Chicago 46, Illinois 1964, page 71.
3. William H. Enneis; Personnel Testing and Equal Employment Opportunity Personnel Testing and Equal Employment Opportunity Commission; U.S. Government Printing Office Washington, D. C. 20402 1970, page 25.
4. Equal Employment Opportunity Commission. Guidelines on Employment Testing Procedures. Washington; EEDC, 1966.
5. Equal Employment Opportunity Commission. part-1607-Guidelines on Employee Selection Procedures; Federal Register Volume 35, Number 149 Saturday, Aug 1, 1970. Washington, D.C. pages 12333-12336.
6. Supreme Court of The United States. Griggs et. al.versus Duke Power Company. No. 124-October Term, 1970 (March 8, 1971).
7. Subcommittee on Labor of the Committee on Labor and Public Welfare United States Senate; The Equal Employment Opportunity Act of 1972; U.S. Government Printing Office; Washington 1972. page 16.
8. American Psychological Association, American Educational Research Association, and National council on Measurement in Education (Joint Committee), Standards for Educational and Psychological Tests; American Psychological Association Inc. 1200 Seventeenth Street, N. W. Washington, D.C. 20036; 1974, page 1.
9. Guion, Robert M. Open a New Window: Validities and Values in Psychological Measurement; American Psychologist Vol. 29, number 5; The American Psychological Association, Inc. May 1974, page 290.
10. Cleary, Anne T. Test bias: Prediction of Grades of Negro and White Students in Integrated Colleges. Journal of Educational Measurement, Vol. 5, No2, Summer 1968, pages 115-124.
11. Thorndike, Robert L. Concepts of Culture-Fairness, Journal of Educational Measurement Vol. 8, No. 2, Summer 1971. pages 63-70.
12. Darlington, Richard B. Another Look at "Cultural Fairness", Journal of Educational Measurement, Vol. 8, No. 2, Summer 1971, pages 71-82.

OTHER REFERENCES

American Educational Research Association, and the National Council on Measurement in Education. "Technical Recommendations for Achievement Tests" 1955.

American Psychological Association; American Educational Research Association, and National Council on Measurements Used in Education (Joint Committee). Technical Recommendations for Psychological tests and diagnostic techniques. Psychological Bulletin, 1954, 51,201-238.

American Psychological Association."Ethical Standards of Psychologists." Washington, D.C., 1953.

Executive Order 11246. Equal Employment Opportunity, Federal Register, Sept, 28, 1965. 12319.

Executive Order 11375. Amending Executive Order No. 11246, Relating to Equal Employment Opportunity. Federal Register, Oct, 17, 1967, 14303.

Office of Federal Contract Compliance, Equal Employment Opportunity, Department of Labor; Part 60-3. Employee Testing and Other Procedures. Federal Register, 1971, 36 (77), 7532-7535.